A

NotI
```
GCGGCCGCCACC ATG AAG CGA TGG AAA AAG AAT TTC ATA GCC GTC TCA GCA GCC AAC CGC TTT AAG AAA ATC TCA TCC TCC
             M   K   R   W   K   K   N   F   I   A   V   S   A   A   N   R   F   K   K   I   S   S   S
```
CBP tag

Gly/Ser-linker
```
GGG GCA CTT GGA AGC GGT AGC GGT ACC ATG GAG GAG AAG ACC ACC GGG TGG CGG GGC GTG GAG GGC CTG GCC
G   A   L   G   S   G   S   G   S   M   D   E   K   T   T   G   W   R   G   G   H   V   E   G   L   A
```
SBP tag (SB19C4)

```
GGC GAG CAG CTC GAG CTG CGG AGG CAC CCT CAG GGC GAG CCC TCC GGC GGC TGC AAG CTG GGC
G   E   L   E   Q   L   R   A   R   L   E   H   P   Q   G   Q   R   E   P   S   G   G   C   K   L   G
```

* GCCCGGGGCGGATCC

tag length: 78AA

B

XhoI   G/S-linker
```
CTC GAG GGA AGC GGT AGC GGT ACC ATG GAC GAG AAG ACC ACC GGG TGG CGG GGC GTG GAG GGC CTG GCC GGC
L   E   G   S   G   S   G   S   M   D   E   K   T   T   G   W   R   G   G   H   V   E   G   L   A   G
```
SBP-tag

```
GAG CAG CTG GAG CTG CGG AGG GCC AGG CTG GAG CAC CCT CAG GGG CAG CGG GAG CCC TCC GGC GGC TGC AAG CTG GGC TCC
E   Q   L   E   L   R   A   R   L   E   H   P   Q   G   Q   R   E   P   S   G   G   C   K   L   G   S
```

```
GGA AAG CGA CGA TGG AAA AAG AAT TTC ATA GCC GTC TCA GCA GCC AAC CGC TTT AAG AAA ATC TCA TCC TCC GGG GCA CTT
G   K   R   R   W   K   K   N   F   I   A   V   S   A   A   N   R   F   K   K   I   S   S   S   G   A   L
```
CBP-tag

ApaI
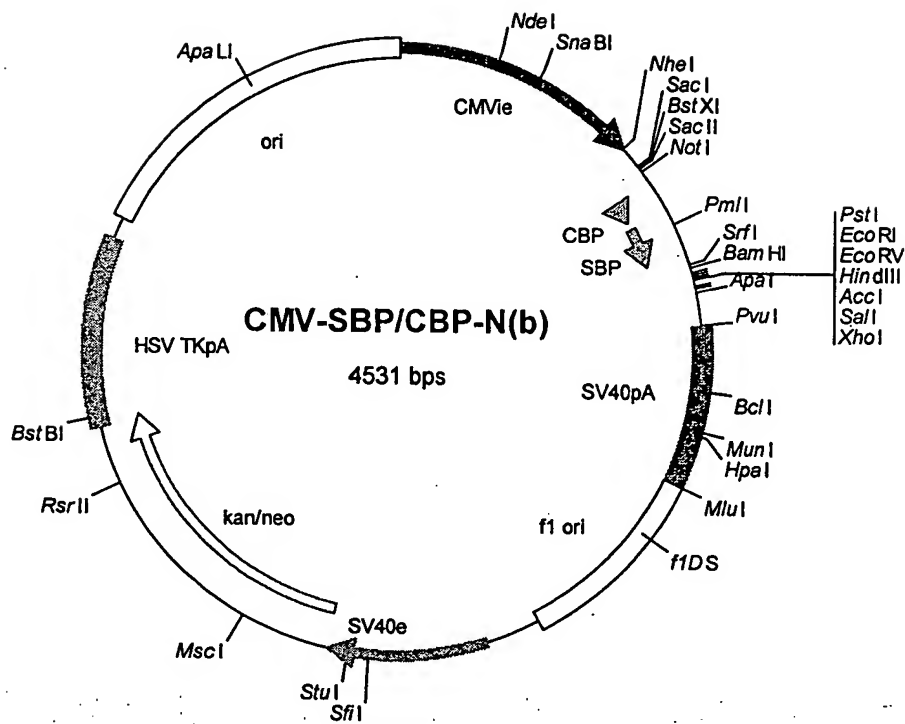TAG GGCCCGAC

Tag length: 79 AA

Fig. 1

Name   No.   Sequence

SB1   3   MDEKTHCTISMNGAVPLVPHHPQGDPLRLLHRPQPALLVRHPQGDLVALVEHHEGVDRGLVALPELHAEELGEPVGDLVQGPVEQVQGVVDALVWRLPPS

SB2   2   MDEKTHCFHPGDHLVRLVEELQALAEGLQROGGROPHRLPRRRPHHLQLLLDEAHPQAGPLREAHQVDGRLLQHHPQGDRLLQQPQDHPLELVWRLPPS

SB3   4   MTRRPTASSSSCVRHLLLRQGEHGHQALEDRDKARHVRLVEGDVEVLGGLDRLARARHEALHPQAGLVHLPLHGGDLGGHLRLVLEAHPQGDRLGLAVHHH

SB4   1   MDEKTHWGISTWRGEPLLHEPQAGRLPLDRRRARHRRILGABPGGVDHGLRLRLELLDDHRPLVPDHHPQRGPLQRGDLPQVVPLVRLRRAHVLGLAAATIT

SB5   3   MDEKTHWVNVHHPQGDLLVRGHGHDVEALHDQGLHQLDLLVGPPEVVRALRGEVLGGLRRIVPLDHPQGEALDQARQRPQHLLELHHRALPPALVWRLPPS

SB6   1   MDEKTHWLNNFEELLARLDGLREGEDHPLVLRHHPQGDGLLDQPLGRHRRALDGEVREGDRPLDQGGEEDLGALVDDDGEVLDGLVHVGVHVHDPLVCGCHHH

SB7   1   MDEKTHWFGTLNSPPTHWMSAVGNGKIDCSFNMNLSLNHWLSSGHPDGALDDQLHPQGDALVGRDDGVVQALRLEGOHQHRRLAQRADRHRQLVWRLPPS

SB8   1   MDEKTHCTIELNFSFTHWKLHHHPQGDALLDDGVRPHHPLADEGGGLDQGLGHRRGVVAERLARRDPEVLEGLVERHRGLVPRLRHGGERHAEPLVWRLPPS

SB9   1   MDEKTHCNTGLYDGAADCFNELNKDVAPLVEGRHDLVEGLLLERHPQGDPLVAHRQLVHPLLGRGERHRRALVPQQEHQPHRLQPVVDLGRRRLVWRLPPS

SB10   1   MDEKTHWHBRAQELVGGILLHDEPQRLLLEPRGPRPLRGLVHEERGHQPQPLAGRVEEADGGLLRDGGELEPLVREGEDHLEPLDDELDAGRGLVWRLPHHH

SB11   1   MDEKTHWHERVHHLADGLLEQHPQGQRRPLVERRHRQVPRGLVRELQHEGLPLRHPAGVHVTRLHQGDRDVDGLVDGHGRDVRGLEREVDGPHRLVWRLPPS

SB12   4   MDKDPLLEELEELRERLVHHPQGGILLPLRGQVGHDABRLGAEVDDLRGGLLDEPQRAVAGLHHVPHRVGQRLVHEVRELDEGLLDQRDDLRQRLVWRLPPS

SB13   2   MEREDPLDEQLRELRELRREALVDHPQGGAQALHRHDGGEHVPLRRVQHRLQPGLLIGELQARLQPLAGEHEGDGAGLQRVPGHQGRRLVWRLPPS

SB14   1   MDEKTHRTLSVSLSFNDWLGQTKACWRLVEGLHGHPQGLVREHEVDVLPLAEEVQQVVGGLADGVSQPGGGLLHRAQRVDHPLPDHAGQVLGRLVWRLPPS

SB15   1   MDEKTHWLEDLKGVLKDCLKDLMDFTKDCRSPRVQPQPLLHHDRGEPVPLLREAGRDLGGLGPRAPRQARPLHHGRHDLHEPIVLQDEPQGGPLVCGCHHH

SB16   1   MDEKTHWVLQLHPQGDRLGPRHGGDDVRLVGQGEGVLEGLLDGRPRRRRRLPREDEHRVRALVDQVRDLAERLVEEVDGGVEALRHLGLPQDEPRSGGCHHH

SB17   2   MDEKTHWGDLQEPLGPLHGGVGEVPGGLVLRHHPQRDRLVDGVGPHGRALARRPHRVVEGLHHLLQRGGERLPPDGPRQLGLLGGELDRADPALVWRLPPS

SB18   1   MDEKTHCAVNVNVGLTHWCHRVAHLQPLDPHPQGDHLRLEPLGHALVDPLVQGVEEVVRPLQLDVGVQRVALVEQVABVGEGLDHEAGQAHGALVWRLPPS

SB19   1   MDEKTGWRGGHVVEGLAGELEQLRARLEHHPQGQREPLVQEVDVDEGLVAGLLDHGVVAGLLDPVEKLLTDWFKKFKNVSKDCKMTFYLEMYDWSGGCHHH

SB20   1   MNEKTHCKLNFKVNIADWLAEFHGGGQGLLGRRDGVVQRLVDGVQERVERLDRDPGLSGDLRLELHHRDHRLRLGGEHLLRDHPLEPDDHLVVGGLVWRLPPS

The "No." indicates the number of times each sequence was observed. The HPQ sequence is in bold type. Defined sequences at the termini are underlined.
The six C-terminal residues are not shown.

Fig. 2

Fig. 3a-1


```
   1  atgcattagt tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga
  61  gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca acgacccccg
 121  cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggа ctttccattg
 181  acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca
 241  tatgccaagt acgccccta ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc
 301  ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat tagtcatcgc
 361  tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc
 421  acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa
 481  tcaacgggac tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag
 541  gcgtgtacgg tgggaggtct atataagcag agctggttta gtgaaccgtc agatccgcta
 601  gcgattacgc caagctcgaa attaaccctc actaaaggga acaaaagctg gagctccacc
 661  gcggtggcgg ccgccaccat gaagcgacga tggaaaaaga atttcatagc cgtctcagca
 721  gccaaccgct ttaagaaaat ctcatcctcc ggggcacttg aagcggtag cggtaccatg
 781  gacgagaaga ccaccggctg gcggggcggc cacgtggtgg agggcctggc cggcgagctg
 841  gagcagctgc gggccaggct ggagcaccac cctcagggcc agcgggagcc ctccggcggc
 901  tgcaagctgg gctgcccggg cggatccccc gggctgcagg aattcgatat caagcttatc
 961  gataccgtcg acctcgaggg ggggcccggt accttaatta attaaggtac caggtaagtg
1021  tacccaattc gccctatagt gagtcgtatt acaattcact cgatcgccct tcccaacagt
1081  tgcgcagcct gaatggcgaa tggagatcca atttttaagt gtataatgtg ttaaactact
1141  gattctaatt gtttgtgtat tttagattca cagtcccaag gctcatttca ggcccctcag
1201  tcctcacagt ctgttcatga tcataatcag ccataccaca tttgtagagg ttttacttgc
```
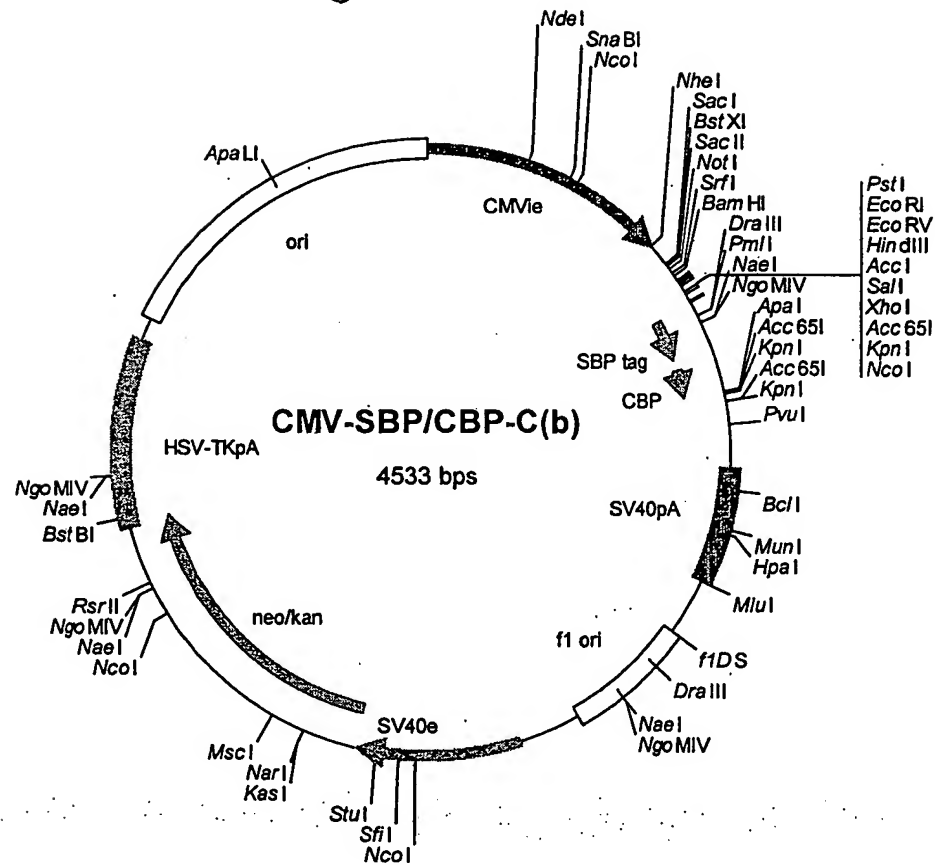
3(a) -1

```
1261  tttaaaaaac ctcccacacc tcccctgaa cctgaaacat aaaatgaatg caattgttgt
1321  tgttaacttg tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt
1381  cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt
1441  atcttaacgc gtaaattgta agcgttaata ttttgttaaa attcgcgtta aattttgtt
1501  aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag
1561  aatagaccga gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga
1621  acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg
1681  aaccatcacc ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc
1741  ctaaagggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg cgagaaagg
1801  aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc
1861  gcgtaaccac cacacccgcc gcgcttaatg cgccgctaca gggcgcgtca ggtggcactt
1921  ttcggggaaa tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaatatgt
1981  atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagaatc
2041  ctgaggcgga aagaaccagc tgtggaatgt gtgtcagtta gggtgtggaa agtccccagg
2101  ctccccagca ggcagaagta tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg
2161  aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc
2221  aaccatagtc ccgcccctaa ctccgcccat cccgcccota actccgccca gttccgccca
2281  ttctccgccc catggctgac taatttttt tatttatgca gaggccgagg ccgcctcggc
2341  ctctgagcta ttccagaagt agtgaggagg cttttttgga ggcctaggct tttgcaaaga
2401  tcgatcaaga gacaggatga ggatcgtttc gcatgattga acaagatgga ttgcacgcag
2461  gttctccggc cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg
2521  gctgctctga tgccgccgtg ttccggctgt cagcgcaggg gcgcccggtt ctttttgtca
2581  agaccgacct gtccggtgcc ctgaatgaac tgcaagacga ggcagcgcgg ctatcgtggc
2641  tggccacgac gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg
2701  actggctgct attgggcgaa gtgccggggc aggatctcct gtcatctcac cttgctcctg
2761  ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta
2821  cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag
2881  ccggtcttgt cgatcaggat gatctggacg aagaacatca ggggctcgcg ccagccgaac
2941  tgttcgccag gctcaaggcg agcatgcccg acggcgagga tctcgtcgtg acccatggcg
3001  atgcctgctt gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg
3061  gccggctggg tgtggcggac cgctatcagg acatagcgtt ggctacccgt gatattgctg
3121  aagaacttgg cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg
3181  attcgcagcg catcgccttc tatcgccttc ttgacgagtt cttctgagcg ggactctggg
3241  gttcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttcg attccaccgc
3301  cgccttctat gaaaggttgg gcttcggaat cgttttccgg gacgccggct ggatgatcct
3361  ccagcgcggg gatctcatgc tggagttctt cgcccaccct aggggaggc taactgaaac
3421  acggaaggag acataccgg aaggaacccg cgctatgacg gcaataaaaa gacagaataa
3481  aacgcacggt gttgggtcgt ttgttcataa acgcggggtt cggtcccagg ctggcactc
3541  tgtcgatacc ccaccgagac cccattgggg ccaatacgcc cgcgtttctt ccttttcccc
3601  accccacccc ccaagttcgg gtgaaggccc agggctcgca gccaacgtcg gggcggcagg
3661  ccctgccata gcctcaggtt actcatatat actttagatt gatttaaaac ttcattttta
3721  atttaaaagg atctaggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg
3781  tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga
3841  tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt
3901  ggtttgtttg ccggatcaag agctaccaac tcttttttccg aaggtaactg gcttcagcag
3961  agcgcagata ccaaatactg tccttctagt gtagccgtag ttaggccacc acttcaagaa
4021  ctctgtagca ccgcctacat acctcgctct gctaatcctg ttaccagtgg ctgctgccag
4081  tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca
4141  gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac
4201  cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aagggagaaa
4261  ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc
4321  agggggaaac gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg
4381  tcgattttg tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc
```

Fig. 3a-2

```
4441   cttttttacgg ttcctggcct tttgctggcc ttttgctcac atgttctttc ctgcgttatc
4501   ccctgattct gtggataacc gtattaccgc c
```

Fig. 3a-3

Fig. 3b-1

CMV-SBP/CBP-C(b)
4533 bps

```
   1  atgcattagt  tattaatagt  aatcaattac  ggggtcatta  gttcatagcc  catatatgga
  61  gttccgcgtt  acataactta  cggtaaatgg  cccgcctggc  tgaccgccca  acgacccccg
 121  cccattgacg  tcaataatga  cgtatgttcc  catagtaacg  ccaataggga  ctttccattg
 181  acgtcaatgg  gtggagtatt  tacggtaaac  tgcccacttg  gcagtacatc  aagtgtatca
 241  tatgccaagt  acgccccta   ttgacgtcaa  tgacggtaaa  tggcccgcct  ggcattatgc
 301  ccagtacatg  accttatggg  actttcctac  ttggcagtac  atctacgtat  tagtcatcgc
 361  tattaccatg  gtgatgcggt  tttggcagta  catcaatggg  cgtggatagc  ggtttgactc
 421  acggggattt  ccaagtctcc  accccattga  cgtcaatggg  agtttgtttt  ggcaccaaaa
 481  tcaacgggac  tttccaaaat  gtcgtaacaa  ctccgcccca  ttgacgcaaa  tgggcggtag
 541  gcgtgtacgg  tgggaggtct  atataagcag  agctggttta  gtgaaccgtc  agatccgcta
 601  gcgattacgc  caagctcgaa  attaaccctc  actaaaggga  acaaaagctg  gagctccacc
 661  gcggtggcgg  ccgctctagc  ccgggcggat  cccccgggct  gcaggaattc  gatatcaagc
 721  ttatcgatac  cgtcgacact  cgagggaagc  ggtagcggta  ccatggacga  agaccacc
 781  ggctggcggg  gcggccacgt  ggtggagggc  ctggccggcg  agctggagca  gctgcgggcc
 841  aggctggagc  accaccctca  gggccagcgg  gagccctccg  gcggctgcaa  gctgggctcc
 901  ggaaagcgac  gatggaaaaa  gaatttcata  gccgtctcag  cagccaaccg  ctttaagaaa
 961  atctcatcct  ccggggcact  ttagggcccg  gtaccttaat  taattaaggt  accaggtaag
1021  tgtacccaat  tcgccctata  gtgagtcgta  ttacaattca  ctcgatcgcc  cttcccaaca
1081  gttgcgcagc  ctgaatggcg  aatggagatc  caatttttaa  gtgtataatg  tgttaaacta
1141  ctgattctaa  ttgtttgtgt  attttagatt  cacagtccca  aggctcattt  caggcccctc
```
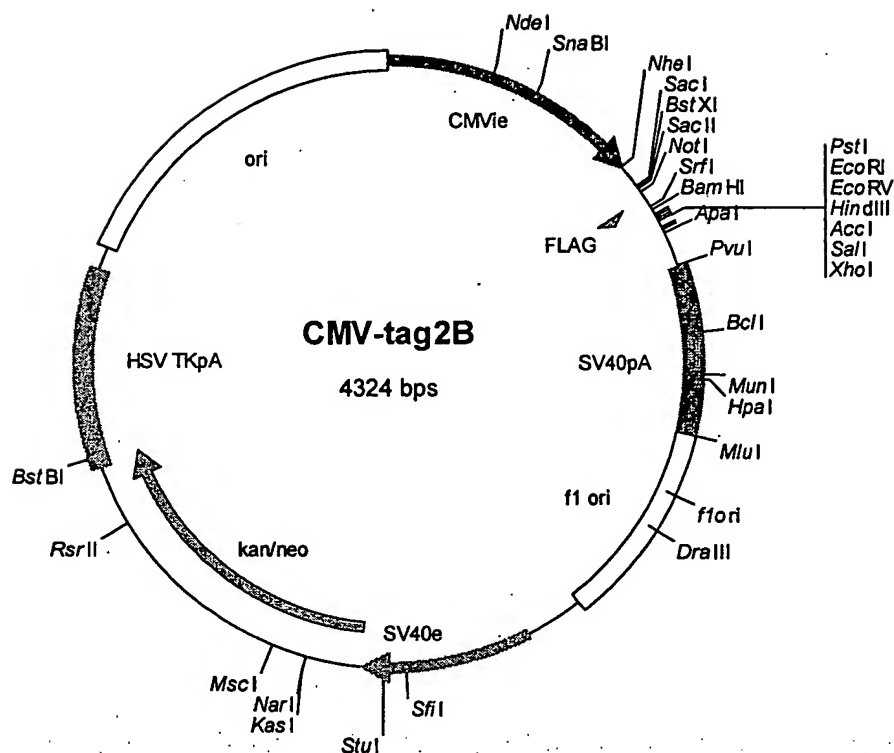
```
1201  agtcctcaca gtctgttcat gatcataatc agccatacca catttgtaga ggttttactt
1261  gctttaaaaa acctcccaca cctccccctg aacctgaaac ataaaatgaa tgcaattgtt
1321  gttgttaact tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaaat
1381  ttcacaaata aagcattttt ttcactgcat tctagttgtg gtttgtccaa actcatcaat
1441  gtatcttaac gcgtaaattg taagcgttaa tattttgtta aaattcgcgt taaattttttg
1501  ttaaatcagc tcattttttta accaataggc cgaaatcggc aaaatcccttt ataaatcaaa
1561  agaatagacc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa
1621  gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg
1681  tgaaccatca ccctaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa
1741  ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa
1801  ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg gcaagtgtag cggtcacgct
1861  gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta cagggcgcgt caggtggcac
1921  ttttcgggga aatgtgcgcg gaacccctat ttgtttattt ttctaaatac attcaaatat
1981  gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa aaaggaagaa
2041  tcctgaggcg gaaagaacca gctgtggaat gtgtgtcagt tagggtgtgg aaagtcccca
2101  ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc aaccaggtgt
2161  ggaaagtccc caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca
2221  gcaaccatag tcccgcccct aactccgccc atcccgcccc taactccgcc cagttccgcc
2281  cattctccgc cccatggctg actaattttt tttatttatg cagaggccga ggccgcctcg
2341  gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg cttttgcaaa
2401  gatcgatcaa gagacaggat gaggatcgtt tcgcatgatt gaacaagatg gattgcacgc
2461  aggttctccg gccgcttggg tggagaggct attcggctat gactgggcac aacagacaat
2521  cggctgctct gatgccgccg tgttccggct gtcagcgcag gggcgcccgg ttcttttttgt
2581  caagaccgac ctgtccggtg ccctgaatga actgcaagac gaggcagcgc ggctatcgtg
2641  gctggccacg acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag
2701  ggactggctg ctattgggcg aagtgccggg gcaggatctc ctgtcatctc accttgctcc
2761  tgccgagaaa gtatccatca tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc
2821  tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta ctcggatgga
2881  agccggtctt gtcgatcagg atgatctgga cgaagaacat caggggctcg cgccagccga
2941  actgttcgcc aggctcaagg cgagcatgcc cgacggcgag gatctcgtcg tgacccatgg
3001  cgatgcctgc ttgccgaata tcatggtgga aaatggccgc ttttctggat tcatcgactg
3061  tggccggctg ggtgtggcgg accgctatca ggacatagcg ttggctaccc gtgatattgc
3121  tgaagaactt ggcggcgaat gggctgaccg cttcctcgtg ctttacggta tcgccgctcc
3181  cgattcgcag cgcatcgcct tctatcgcct tcttgacgag ttcttctgag cgggactctg
3241  gggttcgaaa tgaccgacca agcgacgccc aacctgccat cacgagattt cgattccacc
3301  gccgccttct atgaaaggtt gggcttcgga atcgttttcc gggacgccgg ctggatgatc
3361  ctccagcgcg gggatctcat gctggagttc ttcgcccacc ctagggggag ctaactgaa
3421  acacggaagg agacaatacc ggaaggaacc gcgctatga cggcaataaa aagacagaat
3481  aaaacgcacg gtgttgggtc gtttgttcat aaacgcgggg ttcggtccca gggctggcac
3541  tctgtcgata ccccaccgag accccattgg gccaatacg cccgcgtttc ttccttttcc
3601  ccaccccacc ccccaagttc gggtgaaggc ccagggctcg cagccaacgt cggggcggca
3661  ggccctgcca tagcctcagg ttactcatat atactttaga ttgatttaaa acttcatttt
3721  taatttaaaa ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa
3781  cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga
3841  gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg
3901  gtggtttgtt tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc
3961  agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag
4021  aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc
4081  agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg
4141  cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac
4201  accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga
4261  aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt
4321  ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag
```

Fig. 3b-2

```
4381   cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg
4441   gcctttttac ggttcctggc cttttgctgg ccttttgctc acatgttctt tcctgcgtta
4501   tcccctgatt ctgtggataa ccgtattacc gcc
```

Fig. 3b-3

Plasmid map labels:

Ndel, SnaBI, Nhel, SacI, BstXI, SacII, NotI, SrfI, BamHI, ApaI

PstI, EcoRI, EcoRV, HindIII, AccI, SalI, XhoI

CMVie

FLAG

ori

Pvul

BclI

MunI, HpaI

MluI

CMV-tag2B
4324 bps

SV40pA

f1 ori

f1ori

DraIII

HSV TKpA

BstBI

RsrII

kan/neo
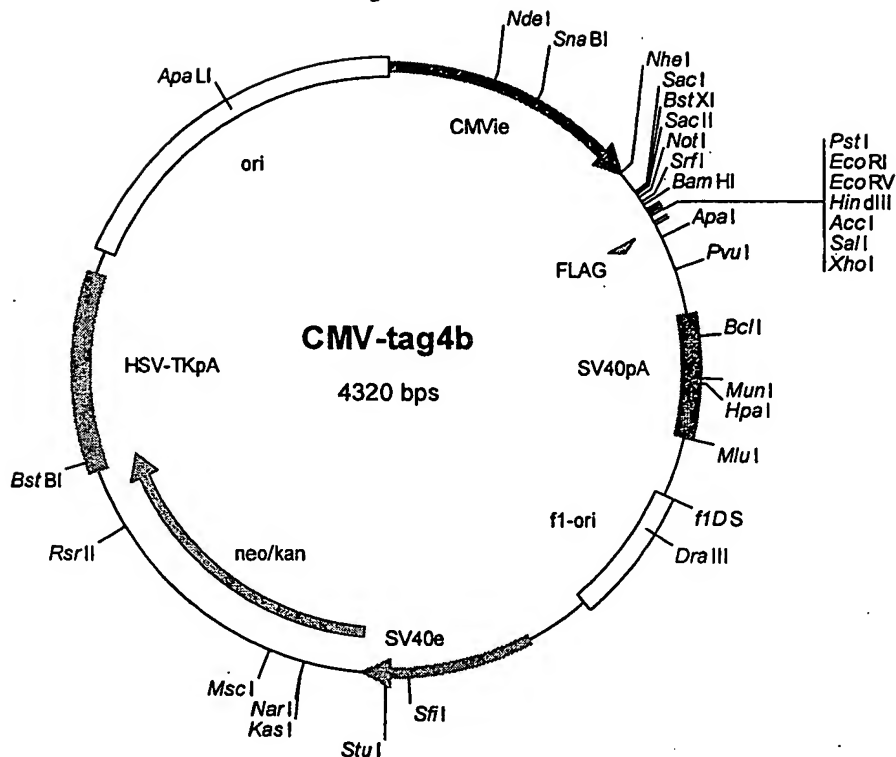
SV40e

MscI, NarI, KasI

StuI

SfiI

```
   1 atgcattagt tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga
  61 gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca acgacccccg
 121 cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggа ctttccattg
 181 acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca
 241 tatgccaagt acgccccta ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc
 301 ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat tagtcatcgc
 361 tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc
 421 acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa
 481 tcaacgggac tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag
 541 gcgtgtacgg tgggaggtct atataagcag agctggttta gtgaaccgtc agatccgcta
 601 gcgattacgc caagctcgaa attaaccctc actaaaggga acaaaagctg gagctccacc
 661 gcggtggcgg ccgccaccat ggattacaag gatgacgacg ataagagccc gggcggatcc
 721 cccgggctgc aggaattcga tatcaagctt atcgataccg tcgacctcga gggggggccc
 781 ggtaccttaa ttaattaagg taccaggtaa gtgtacccaa ttcgccctat agtgagtcgt
 841 attacaattc actcgatcgc ccttcccaac agttgcgcag cctgaatggc gaatggagat
 901 ccaattttta agtgtataat gtgttaaact actgattcta attgtttgtg tattttagat
 961 tcacagtccc aaggctcatt tcaggcccct cagtcctcac agtctgttca tgatcataat
1021 cagccatacc acatttgtag aggttttact tgctttaaaa aacctcccac acctcccct
1081 gaacctgaaa cataaaatga atgcaattgt tgttgttaac ttgtttattg cagcttataa
1141 tggttacaaa taaagcaata gcatcacaaa tttcacaaat aaagcatttt tttcactgca
1201 ttctagttgt ggtttgtcca aactcatcaa tgtatcttaa cgcgtaaatt gtaagcgtta
```

```
1261   atattttgtt aaaattcgcg ttaaattttt gttaaatcag ctcatttttt aaccaatagg
1321   ccgaaatcgg caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg
1381   ttccagtttg gaacaagagt ccactattaa agaacgtgga ctccaacgtc aaagggcgaa
1441   aaaccgtcta tcaggcgat ggcccactac gtgaaccatc accctaatca agttttttgg
1501   ggtcgaggtg ccgtaaagca ctaaatcgga accctaaagg gagcccccga tttagagctt
1561   gacggggaaa gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa ggagcgggcg
1621   ctagggcgct ggcaagtgta gcggtcacgc tgcgcgtaac caccacaccc gccgcgctta
1681   atgcgccgct acagggcgcg tcaggtggca cttttcgggg aaatgtgcgc ggaacccta
1741   tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat
1801   aaatgcttca ataatattga aaaggaaga atcctgaggc ggaaagaacc agctgtggaa
1861   tgtgtgtcag ttagggtgtg aaagtcccc aggctcccca gcaggcagaa gtatgcaaag
1921   catgcatctc aattagtcag caaccaggtg tggaaagtcc ccaggctccc cagcaggcag
1981   aagtatgcaa agcatgcatc tcaattagtc agcaaccata gtcccgcccc taactccgcc
2041   catcccgccc ctaactccgc ccagttccgc ccattctccg ccccatggct gactaatttt
2101   ttttatttat gcagaggccg aggccgcctc ggcctctgag ctattccaga agtagtgagg
2161   aggcttttt ggaggcctag gcttttgcaa agatcgatca agagacagga tgaggatcgt
2221   ttcgcatgat tgaacaagat ggattgcacg caggttctcc ggccgcttgg gtggagaggc
2281   tattcggcta tgactgggca caacagacaa tcggctgctc tgatgccgcc gtgttccggc
2341   tgtcagcgca ggggcgcccg gttctttttg tcaagaccga cctgtccggt gccctgaatg
2401   aactgcaaga cgaggcagcg cggctatcgt ggctggccac gacgggcgtt ccttgcgcag
2461   ctgtgctcga cgttgtcact gaagcgggaa gggactggct gctattgggc gaagtgccgg
2521   ggcaggatct cctgtcatct caccttgctc ctgccgagaa agtatccatc atggctgatg
2581   caatgcggcg gctgcatacg cttgatccgg ctacctgccc attcgaccac caagcgaaac
2641   atcgcatcga gcgagcacgt actcggatgg aagccggtct tgtcgatcag gatgatctgg
2701   acgaagaaca tcaggggctc gcgccagccg aactgttcgc caggctcaag gcgagcatgc
2761   ccgacggcga ggatctcgtc gtgacccatg gcgatgcctg cttgccgaat atcatggtgg
2821   aaaatggccg cttttctgga ttcatcgact gtggccggct gggtgtggcg gaccgctatc
2881   aggacatagc gttggctacc cgtgatattg ctgaagaact ggcggcgaa tgggctgacc
2941   gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcatcgcc ttctatcgcc
3001   ttcttgacga gttcttctga gcgggactct ggggttcgaa atgaccgacc aagcgacgcc
3061   caacctgcca tcacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg
3121   aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt
3181   cttcgcccac cctaggggga ggctaactga aacacggaag gagacaatac cggaaggaac
3241   ccgcgctatg acggcaataa aaagacagaa taaaacgcac ggtgttgggt cgtttgttca
3301   taaacgcggg gttcggtccc agggctggca ctctgtcgat accccaccga gaccccattg
3361   gggccaatac gcccgcgttt cttccttttc cccacccccac ccccaagtt cgggtgaagg
3421   cccagggctc gcagccaacg tcggggcggc aggccctgcc atagcctcag gttactcata
3481   tatactttag attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct
3541   ttttgataat ctcatgacca aaatccctta acgtgagttt cgttccact gagcgtcaga
3601   ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg
3661   cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc
3721   aactcttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct
3781   agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc
3841   tctgctaatc ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt
3901   ggactcaaga cgatagttac cggataaggc gcagcggtcg gctgaacgg ggggttcgtg
3961   cacacagccc agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct
4021   atgagaaagc gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag
4081   ggtcggaaca ggagagcgca cgagggagct ccaggggga aacgcctggt atctttatag
4141   tcctgtcggg tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg
4201   gcggagccta tggaaaaacg ccagcaacgc ggcctttta cggttcctgg ccttttgctg
4261   gccttttgct cacatgttct ttcctgcgtt atccctgat tctgtggata accgtattac
4321   cgcc
```

Fig. 4a-2

Plasmid map labeled CMV-tag4b, 4320 bps, with restriction sites ApaLI, NdeI, SnaBI, NheI, SacI, BstXI, SacII, NotI, SrfI, BamHI, PstI, EcoRI, EcoRV, HindIII, AccI, SalI, XhoI, ApaI, PvuI, BclI, MunI, HpaI, MluI, DraIII, f1DS, SfiI, StuI, KasI, NarI, MscI, RsrII, BstBI, and features CMVie, FLAG, SV40pA, f1-ori, SV40e, neo/kan, HSV-TKpA, ori.

```
   1 atgcattagt tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga
  61 gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca acgacccccg
 121 cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatatggg actttccattg
 181 acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca
 241 tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc
 301 ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat tagtcatcgc
 361 tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc
 421 acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa
 481 tcaacgggac tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag
 541 gcgtgtacgg tgggaggtct atataagcag agctggttta gtgaaccgtc agatccgcta
 601 gcgattacgc caagctcgaa attaaccctc actaaaggga acaaaagctg gagctccacc
 661 gcggtggcgg ccgctctagc ccgggcggat cccccgggct gcaggaattc gatatcaagc
 721 ttatcgatac cgtcgacact cgaggattac aaggatgacg acgataagta gggcccggta
 781 ccttaattaa ttaaggtacc aggtaagtgt acccaattcg ccctatagtg agtcgtatta
 841 caattcactc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggagatccaa
 901 tttttaagtg tataatgtgt taaactactg attctaattg tttgtgtatt ttagattcac
 961 agtcccaagg ctcatttcag gcccctcagt cctcacagtc tgttcatgat cataatcagc
1021 cataccacat ttgtagaggt tttacttgct ttaaaaaacc tcccacacct ccccctgaac
1081 ctgaaacata aaatgaatgc aattgttgtt gttaacttgt ttattgcagc ttataatggt
1141 tacaaataaa gcaatagcat cacaaatttc acaaataaag catttttttc actgcattct
1201 agttgtggtt tgtccaaact catcaatgta tcttaacgcg taaattgtaa gcgttaatat
1261 tttgttaaaa ttcgcgttaa attttttgtta aatcagctca ttttttaacc aataggccga
1321 aatcggcaaa atcccttata aatcaaaaga atagaccgag ataggttga gtgttgttcc
1381 agtttggaac aagagtccac tattaaagaa cgtggactcc aacgtcaaag ggcgaaaaac
1441 cgtctatcag ggcgatggcc cactacgtga accatcaccc taatcaagtt ttttggggtc
1501 gaggtgccgt aaagcactaa atcggaaccc taaagggagc ccccgattta gagcttgacg
1561 gggaaagccg gcgaacgtgg cgagaaagga agggaagaaa gcgaaaggag cgggcgctag
```

```
1621  ggcgctggca  agtgtagcgg  tcacgctgcg  cgtaaccacc  acacccgccg  cgcttaatgc
1681  gccgctacag  ggcgcgtcag  gtggcacttt  tcggggaaat  gtgcgcggaa  cccctatttg
1741  tttatttttc  taaatacatt  caaatatgta  tccgctcatg  agacaataac  cctgataaat
1801  gcttcaataa  tattgaaaaa  ggaagaatcc  tgaggcggaa  agaaccagct  gtggaatgtg
1861  tgtcagttag  ggtgtggaaa  gtccccaggc  tccccagcag  gcagaagtat  gcaaagcatg
1921  catctcaatt  agtcagcaac  caggtgtgga  aagtccccag  gctccccagc  aggcagaagt
1981  atgcaaagca  tgcatctcaa  ttagtcagca  accatagtcc  cgcccctaac  tccgcccatc
2041  ccgcccctaa  ctccgcccag  ttccgcccat  tctccgcccc  atggctgact  aatttttttt
2101  atttatgcag  aggccgaggc  cgcctcggcc  tctgagctat  tccagaagta  gtgaggaggc
2161  ttttttggag  gcctaggctt  ttgcaaagat  cgatcaagag  acaggatgag  gatcgtttcg
2221  catgattgaa  caagatggat  tgcacgcagg  ttctccggcc  gcttgggtgg  agaggctatt
2281  cggctatgac  tgggcacaac  agacaatcgg  ctgctctgat  gccgccgtgt  tccggctgtc
2341  agcgcagggg  cgcccggttc  tttttgtcaa  gaccgacctg  tccggtgccc  tgaatgaact
2401  gcaagacgag  gcagcgcggc  tatcgtggct  ggccacgacg  ggcgttcctt  gcgcagctgt
2461  gctcgacgtt  gtcactgaag  cgggaaggga  ctggctgcta  ttgggcgaag  tgccggggca
2521  ggatctcctg  tcatctcacc  ttgctcctgc  cgagaaagta  tccatcatgg  ctgatgcaat
2581  gcggcggctg  catacgcttg  atccggctac  ctgcccattc  gaccaccaag  cgaaacatcg
2641  catcgagcga  gcacgtactc  ggatggaagc  cggtcttgtc  gatcaggatg  atctggacga
2701  agaacatcag  gggctcgcgc  cagccgaact  gttcgccagg  ctcaaggcga  gcatgcccga
2761  cggcgaggat  ctcgtcgtga  cccatggcga  tgcctgcttg  ccgaatatca  tggtggaaaa
2821  tggccgcttt  tctggattca  tcgactgtgg  ccggctgggt  gtggcggacc  gctatcagga
2881  catagcgttg  gctacccgtg  atattgctga  agaacttggc  ggcgaatggg  ctgaccgctt
2941  cctcgtgctt  tacggtatcg  ccgctcccga  ttcgcagcgc  atcgccttct  atcgccttct
3001  tgacgagttc  ttctgagcgg  gactctgggg  ttcgaaatga  ccgaccaagc  gacgcccaac
3061  ctgccatcac  gagatttcga  ttccaccgcc  gccttctatg  aaaggttggg  cttcggaatc
3121  gttttccggg  acgccggctg  gatgatcctc  cagcgcgggg  atctcatgct  ggagttcttc
3181  gcccacccta  ggggggaggct  aactgaaaca  cggaaggaga  caataccgga  aggaacccgc
3241  gctatgacgg  caataaaaag  acagaataaa  acgcacggtg  ttgggtcgtt  tgttcataaa
3301  cgcggggttc  ggtcccaggg  ctggcactct  gtcgataccc  caccgagacc  ccattggggc
3361  caatacgccc  gcgtttcttc  cttttcccca  ccccaccccc  caagttcggg  tgaaggccca
3421  gggctcgcag  ccaacgtcgg  ggcggcaggc  cctgccatag  cctcaggtta  ctcatatata
3481  ctttagattg  atttaaaact  tcatttttaa  tttaaaagga  tctaggtgaa  gatcctttttt
3541  gataatctca  tgaccaaaat  cccttaacgt  gagttttcgt  tccactgagc  gtcagacccc
3601  gtagaaaaga  tcaaaggatc  ttcttgagat  cctttttttc  tgcgcgtaat  ctgctgcttg
3661  caaacaaaaa  aaccaccgct  accagcggtg  gtttgtttgc  cggatcaaga  gctaccaact
3721  cttttttccga  aggtaactgg  cttcagcaga  gcgcagatac  caaatactgt  ccttctagtg
3781  tagccgtagt  taggccacca  cttcaagaac  tctgtagcac  cgcctacata  cctcgctctg
3841  ctaatcctgt  taccagtggc  tgctgccagt  ggcgataagt  cgtgtcttac  cgggttggac
3901  tcaagacgat  agttaccgga  taaggcgcag  cggtcgggct  gaacgggggg  ttcgtgcaca
3961  cagcccagct  tggagcgaac  gacctacacc  gaactgagat  acctacagcg  tgagctatga
4021  gaaagcgcca  cgcttcccga  agggagaaag  gcggacaggt  atccggtaag  cggcagggtc
4081  ggaacaggag  agcgcacgag  ggagcttcca  gggggaaacg  cctggtatct  ttatagtcct
4141  gtcgggtttc  gccacctctg  acttgagcgt  cgatttttgt  gatgctcgtc  aggggggcgg
4201  agcctatgga  aaaacgccag  caacgcggcc  tttttacggt  tcctggcctt  ttgctggcct
4261  tttgctcaca  tgttctttcc  tgcgttatcc  cctgattctg  tggataaccg  tattaccgcc
```
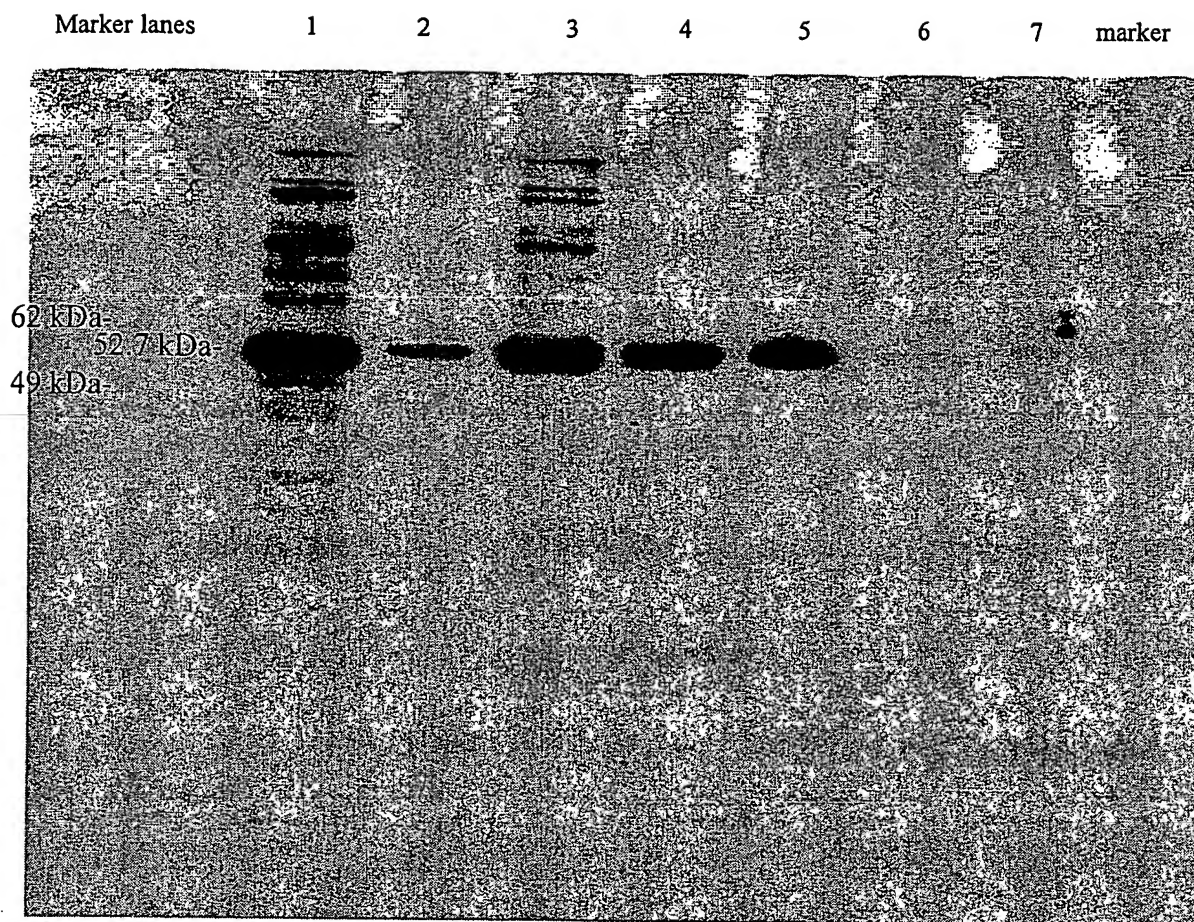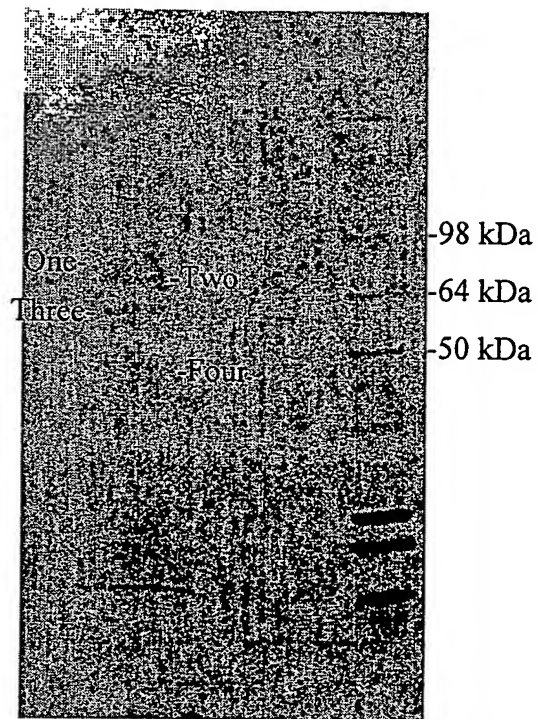
Fig. 4b-2

FIGURE 5

FIGURE 6